

Optimal Local and Remote Controllers with Unreliable Communication

Yi Ouyang, Seyed Mohammad Asghari, and Ashutosh Nayyar

Abstract—We consider a decentralized optimal control problem for a linear plant controlled by two controllers, a local controller and a remote controller. The local controller directly observes the state of the plant and can inform the remote controller of the plant state through a packet-drop channel. We assume that the remote controller is able to send acknowledgments to the local controller to signal the successful receipt of transmitted packets. The objective of the two controllers is to cooperatively minimize a quadratic performance cost. We provide a dynamic program for this decentralized control problem using the common information approach. Although our problem is not a partially nested LQG problem, we obtain explicit optimal strategies for the two controllers. In the optimal strategies, both controllers compute a common estimate of the plant state based on the common information. The remote controller's action is linear in the common estimated state, and the local controller's action is linear in both the actual state and the common estimated state.

I. INTRODUCTION

Networked control systems (NCS) are distributed systems that consist of several components (e.g. physical systems, controllers, smart sensors, etc.) and the communication network that connects them together. With the recent interest in cyber-physical systems and the Internet of Things (IoT), NCS have received considerable attention in the recent years (see [1] and references therein). In contrast to traditional control systems, the interconnected components in NCS are linked through unreliable channels with random packet drops and delays. In the presence of unreliable communication in NCS, the implicit assumption of perfect data exchange in classical estimation and control system fails [2]. Therefore, efficient operation of NCS requires decentralized decision-making while taking into account the unreliable communication among decision-makers.

In this paper, we consider an optimal control problem for a NCS consisting of a linear plant and two controllers, namely the local controller and the remote controller, connected through an unreliable communication link as shown in Fig. 1. The local controller directly observes the state of the plant and can inform the remote controller of the plant state through a channel with random packet drops. We consider a TCP structure so that the remote controller is able to send acknowledgments to the local controller to signal the successful receipt of transmitted packets. The objective of the two controllers is to cooperatively minimize the overall quadratic performance cost of the NCS. The problem is motivated from applications that demand remote control of systems over wireless networks where links are prone to failure. The local controller can be a small local processor proximal to the system that measures the status of the system and can perform limited control. The remote controller can

be a more powerful controller that receives information from the local processor through a wireless channel.

Similar setups of NCS has been investigated in the literature with only the remote controller present. Various communication protocols including the TCP (where acknowledgments are available) and the UDP (where acknowledgments are not available) and variations have been investigated [3–7]. For NCS with two decision-makers, [8], [9] have studied the problem when the local controller is a smart sensor and the remote controller is an estimator. When the linear plant is controlled only by the remote controller and the local controller is a smart sensor or encoder, [10–14] have shown that the separation of control and estimation holds for the remote controller under various communication channel models.

The problem considered in this paper is different from previous works on NCS because our problem is a two-controller decentralized problem where both controllers can control the dynamics of the plant. Finding optimal strategies for two-controller decentralized problems is generally difficult (see [15–17]). In general, linear control strategies are not optimal, and even the problem of finding the best linear control strategies is not convex [18]. Existing optimal solutions of two-controller decentralized problems require either specific information structures, such as static [19], partially nested [20–25], stochastically nested [26], or other specific properties, such as quadratic invariance [27] or substitutability [28]. None of the above properties hold in our problem due to either the unreliable communication or the nature of dynamics and cost function. In spite of this, we solve the two-controller decentralized problem and provide explicit optimal strategies for the local controller and the remote controller. In the optimal strategies, both controllers compute a common estimate of the plant state based on the common information. The remote controller's action is linear in the common estimated state, and the local controller's action is linear in both the actual state and the common estimated state.

A. Organization

The rest of the paper is organized as follows. We introduce the system model and formulate the two-controller optimal control problem in Section II. In Section III, we provide a dynamic program for the decentralized control problem using the common information approach. We solve the dynamic program in Section IV. Section V concludes the paper.

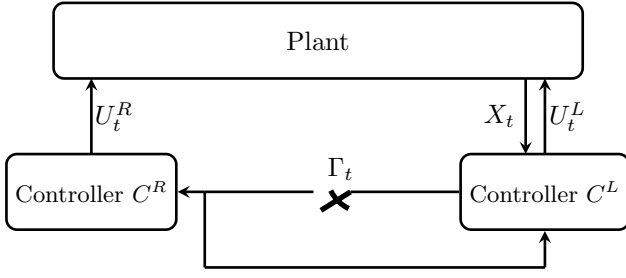


Fig. 1. Two-controller system model. The binary random variable Γ_t indicates whether packets are transmitted successfully.

Notation

Random variables/vectors are denoted by upper case letters, their realization by the corresponding lower case letter. For a sequence of column vectors X, Y, Z, \dots , the notation $\text{vec}(X, Y, Z, \dots)$ denotes vector $[X^\top, Y^\top, Z^\top, \dots]^\top$. The transpose and trace of matrix A are denoted by A^\top and $\text{tr}(A)$, respectively. In general, subscripts are used as time index while superscripts are used to index controllers. For time indices $t_1 \leq t_2$, $X_{t_1:t_2}$ (resp. $g_{t_1:t_2}(\cdot)$) is the short hand notation for the variables $(X_{t_1}, X_{t_1+1}, \dots, X_{t_2})$ (resp. functions $(g_{t_1}(\cdot), \dots, g_{t_2}(\cdot))$). The indicator function of set E is denoted by $\mathbb{1}_E(\cdot)$, that is, $\mathbb{1}_E(x) = 1$ if $x \in E$, and 0 otherwise. $\mathbb{P}(\cdot)$, $\mathbb{E}[\cdot]$, and $\text{cov}(\cdot)$ denote the probability of an event, the expectation of a random variable/vector, and the covariance matrix of a random vector, respectively. For random variables/vectors X and Y , $\mathbb{P}(\cdot|Y = y)$ denotes the probability of an event given that $Y = y$, and $\mathbb{E}[X|y] := \mathbb{E}[X|Y = y]$. For a strategy g , we use $\mathbb{P}^g(\cdot)$ (resp. $\mathbb{E}^g[\cdot]$) to indicate that the probability (resp. expectation) depends on the choice of g . Let $\Delta(\mathbb{R}^n)$ denote the set of all probability measures on \mathbb{R}^n . For any $\theta \in \Delta(\mathbb{R}^n)$, $\theta(E) = \int_{\mathbb{R}^n} \mathbb{1}_E(x) \theta(dx)$ denotes the probability of event E under θ . The mean and the covariance of a distribution $\theta \in \Delta(\mathbb{R}^n)$ are denoted by $\mu(\theta)$ and $\text{cov}(\theta)$, respectively, and are defined as $\mu(\theta) = \int_{\mathbb{R}^n} x \theta(dx)$ and $\text{cov}(\theta) = \int_{\mathbb{R}^n} (x - \mu(\theta))(x - \mu(\theta))^\top \theta(dx)$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider the discrete-time system with two controllers as shown in Fig. 1. The linear plant dynamics are given by

$$X_{t+1} = AX_t + B^L U_t^L + B^R U_t^R + W_t, t = 0, \dots, T \quad (1)$$

where $X_t \in \mathbb{R}^{n_x}$ is the state of the plant at time t , $U_t^L \in \mathbb{R}^{n_L}$ is the control action of the local controller C^L , $U_t^R \in \mathbb{R}^{n_R}$ is the control action of the remote controller C^R , and A, B^L, B^R are matrices with appropriate dimensions. X_0 is a random vector with distribution π_{X_0} , $W_t \in \mathbb{R}^{n_x}$ is a zero mean noise vector at time t with distribution π_{W_t} . $X_0, W_0, W_1, \dots, W_T$ are independent random vectors with finite second moments.

At each time t the local controller C^L perfectly observes the state X_t and sends the observed state to the remote controller C^R through an unreliable channel with packet drop probability p . Let Γ_t be Bernoulli random variable describing

the nature of this channel, that is, $\Gamma_t = 0$ when the link is broken and otherwise, $\Gamma_t = 1$. We assume that Γ_t is independent of all other variables before time t . Furthermore, let Z_t be the channel output, then,

$$\Gamma_t = \begin{cases} 1 & \text{with probability } (1-p), \\ 0 & \text{with probability } p. \end{cases} \quad (2)$$

$$Z_t = \begin{cases} X_t & \text{when } \Gamma_t = 1, \\ \emptyset & \text{when } \Gamma_t = 0. \end{cases} \quad (3)$$

We assume that the channel output Z_t is perfectly observed by C^R . The remote controller sends an acknowledgment when it receives the state. Thus, effectively, Z_t is perfectly observed by C^L as well. The two controllers select their control actions after observing Z_t . We assume that the links from the controllers to the plant are perfect.

Let H_t^L and H_t^R denote the information available to C^L and C^R to make decisions at time t , respectively.¹ Then,

$$H_t^L = \{X_{0:t}, Z_{0:t}, U_{0:t-1}^L, U_{0:t-1}^R\}, \quad H_t^R = \{Z_{0:t}, U_{0:t-1}^R\}. \quad (4)$$

Let \mathcal{H}_t^L and \mathcal{H}_t^R be the spaces of all possible information of C^L and C^R at time t , respectively. Then, C^L and C^R 's actions are selected according to

$$U_t^L = g_t^L(H_t^L), \quad U_t^R = g_t^R(H_t^R), \quad (5)$$

where the control strategies $g_t^L : \mathcal{H}_t^L \mapsto \mathbb{R}^{n_L}$ and $g_t^R : \mathcal{H}_t^R \mapsto \mathbb{R}^{n_R}$ are measurable mappings.

The instantaneous cost $c_t(X_t, U_t^L, U_t^R)$ of the system is a general quadratic function given by

$$c_t(X_t, U_t^L, U_t^R) = S_t^\top R_t S_t, \text{ where } S_t = \text{vec}(X_t, U_t^L, U_t^R), R_t = \begin{bmatrix} R_t^{XX} & R_t^{XL} & R_t^{XR} \\ R_t^{LX} & R_t^{LL} & R_t^{LR} \\ R_t^{RX} & R_t^{RL} & R_t^{RR} \end{bmatrix},$$

and R_t is a symmetric positive definite (PD) matrix.

The performance of strategies $g^L := g_{0:T}^L$ and $g^R := g_{0:T}^R$ is the total expected cost given by

$$J(g^L, g^R) = \mathbb{E}^{g^L, g^R} \left[\sum_{t=0}^T c_t(X_t, U_t^L, U_t^R) \right]. \quad (6)$$

Let \mathcal{G}^L and \mathcal{G}^R denote all possible control strategies of C^L and C^R respectively. The optimal control problem for C^L and C^R is formally defined below.

Problem 1. For the system described by (1)-(6), determine control strategies g^L and g^R that minimize the performance cost of (6).

Problem 1 is a two-controller decentralized optimal control problem. Note that Problem 1 is not a partially nested LQG problem. In particular, the local controller C^L 's action U_{t-1}^L at $t-1$ affects X_t , and consequently, it affects Z_t . Since Z_t is a part of the remote controller C^R 's information H_t^R at t but $H_{t-1}^L \not\subset H_t^R$, the information structure in Problem 1 is

¹ U_{t-1}^R is not directly observed by C^L at time t , but C^L can obtain U_{t-1}^R because $U_{t-1}^R = g_{t-1}^R(H_{t-1}^R)$ and $H_{t-1}^R \subset H_t^L$.

not partially nested. Therefore, linear control strategies are not necessarily optimal for Problem 1.

Our approach to Problem 1 is based on the common information approach [29] for decentralized decision-making. We identify the common belief of the system state for C^L and C^R . The common belief can serve as an information state that leads to a dynamic program for optimal strategies of the two-controller problem.

Remark 1. *The results of [29] are developed only for finite spaces. Therefore, we can not directly apply the results of [29] to Problem 1.*

III. COMMON BELIEF AND DYNAMIC PROGRAM

From (4), H_t^R is the *common information* among the two controllers. Consider fixed strategies $g_{0:t-1}^L, g_{0:t-1}^R$ until time $t-1$. Given any realization $h_t^R \in \mathcal{H}_t^R$ of the common information, we define the common belief $\theta_t \in \Delta(\mathbb{R}^{n_x})$ as the conditional probability distribution of X_t given h_t^R . That is, for any measurable set $E \subset \mathbb{R}^{n_x}$

$$\theta_t(X_t \in E) = \mathbb{P}^{g_{0:t-1}^L, g_{0:t-1}^R}(X_t \in E | h_t^R). \quad (7)$$

Using ideas from the common information approach [29], the common belief θ_t could serve as an information state for decentralized decision-making. We proceed to show that θ_t is indeed an information state that can be used to write a dynamic program for Problem 1.

The following Lemma provides a structural result for C^L .

Lemma 1. *Let $\hat{H}_t^L = \text{vec}(X_t, H_t^R)$, and $\hat{\mathcal{H}}_t^L$ be the space of all possible \hat{H}_t^L . Let $\hat{\mathcal{G}}^L = \{g^L : g^L \text{ is measurable from } \hat{\mathcal{H}}_t^L \text{ to } \mathbb{R}^{n_L}\}$. Then,*

$$\inf_{g^L \in \hat{\mathcal{G}}^L, g^R \in \mathcal{G}^R} J(g^L, g^R) = \inf_{g^L \in \hat{\mathcal{G}}^L, g^R \in \mathcal{G}^R} J(g^L, g^R). \quad (8)$$

From Lemma 1, we only need to consider strategies $g^L \in \hat{\mathcal{G}}^L$ for the local controller C^L . That is, C^L only needs to use $\hat{H}_t^L = \text{vec}(X_t, H_t^R)$ to make the decision at t .

For any strategy $g^L \in \hat{\mathcal{G}}^L$ we provide a representation of g^L using the space \mathcal{Q}^θ defined below.

Definition 1. *For any $\theta \in \Delta(\mathbb{R}^{n_x})$, define a set of mappings*

$$\mathcal{Q}^\theta = \left\{ q : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_L} \text{ measurable}, \int_{\mathbb{R}^{n_x}} q(x) \theta(dx) = 0 \right\}. \quad (9)$$

Lemma 2. *For any strategies $g^L \in \hat{\mathcal{G}}^L$ and $g^R \in \mathcal{G}^R$, let θ_t be the conditional probability distribution defined in (7). Then at any time t there exists $\bar{g}_t^L : \mathcal{H}_t^R \mapsto \mathbb{R}^{n_L}$ and $\tilde{g}_t^L : \mathcal{H}_t^R \mapsto \mathcal{Q}^{\theta_t}$ such that \bar{g}_t^L is measurable and*

$$g_t^L(x_t, h_t^R) = \bar{g}_t^L(h_t^R) + q_t(x_t), \quad q_t = \tilde{g}_t^L(h_t^R). \quad (10)$$

Proof of Lemma 2. Define

$$\bar{g}_t^L(h_t^R) = \mathbb{E}^{g_{0:t-1}^L, g_{0:t-1}^R} [g_t^L(X_t, h_t^R) | h_t^R], \quad (11)$$

$$q_t(\cdot) = \tilde{g}_t^L(h_t^R)(\cdot) = g_t^L(\cdot, h_t^R) - \bar{g}_t^L(h_t^R). \quad (12)$$

Since $g_t^L(x_t, h_t^R)$ is measurable, $\bar{g}_t^L(h_t^R)$ is also measurable. For each $h_t^R \in \mathcal{H}_t^R$, $q_t(\cdot) = \tilde{g}_t^L(h_t^R)(\cdot)$ is a measurable function because $g_t^L(x_t, h_t^R)$ is measurable. Furthermore,

$$\begin{aligned} \int_{\mathbb{R}^{n_x}} q_t(x) \theta_t(dx) &= \int_{\mathbb{R}^{n_x}} g_t^L(x, h_t^R) \theta_t(dx) \\ &\quad - \mathbb{E}^{g_{0:t-1}^L, g_{0:t-1}^R} [g_t^L(X_t, h_t^R) | h_t^R] = 0. \end{aligned}$$

The last equality follows from (7). Therefore, $q_t \in \mathcal{Q}^{\theta_t}$. \square

Note that q_t belongs to \mathcal{Q}^{θ_t} and is itself a function of h_t^R .

From Lemma 2, for any strategies $g^L \in \hat{\mathcal{G}}^L$ and $g^R \in \mathcal{G}^R$ we have a corresponding representation of the strategy g^L of C^L in terms of \bar{g}_t^L and \tilde{g}_t^L .

Using the above representation of C^L 's strategy, we can show that the common belief θ_t is an information state with a sequential update function.

For any $x \in \mathbb{R}^{n_x}$, let $\delta_x \in \Delta(\mathbb{R}^{n_x})$ denote the Dirac delta distribution at point x , that is, for any measurable set $E \subset \mathbb{R}^{n_x}$, $\delta_x(E) = 1$ if $x \in E$, and otherwise $\delta_x(E) = 0$. Define $\varphi : \mathbb{R}^{n_x} \mapsto \Delta(\mathbb{R}^{n_x})$ such that $\varphi(x) = \delta_x$ for any $x \in \mathbb{R}^{n_x}$.

Lemma 3. *For any strategies $g^L \in \hat{\mathcal{G}}^L$ and $g^R \in \mathcal{G}^R$, let $(\bar{g}_t^L, \tilde{g}_t^L)$ be the representation of g^L given by Lemma 2. Then the common beliefs $\{\theta_t, t = 0, 1, \dots, T\}$, defined by (7), can be sequentially updated according to*

$$\theta_0 = \begin{cases} \pi_{x_0} & \text{if } z_0 = \emptyset, \\ \varphi(x_0) & \text{if } z_0 = x_0. \end{cases} \quad (13)$$

$$\theta_{t+1} = \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, z_{t+1}), \quad (14)$$

where u_t^R, \bar{u}_t^L and q_t are functions of the common information h_t^R given by

$$u_t^R = g_t^R(h_t^R), \quad \bar{u}_t^L = \bar{g}_t^L(h_t^R), \quad q_t = \tilde{g}_t^L(h_t^R). \quad (15)$$

Furthermore, $\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, x_{t+1}) = \varphi(x_{t+1})$ and $\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)$ is a distribution on \mathbb{R}^{n_x} such that for any measurable set $E \subset \mathbb{R}^{n_x}$,

$$\begin{aligned} \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)(E) &= \\ \int_{\mathbb{R}^{n_x}} \int_{\mathbb{R}^{n_x}} \mathbb{1}_E(Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t) \\ &\quad \theta_t(dx_t) \pi_{W_t}(dw_t). \end{aligned} \quad (16)$$

Using the common belief and its update function, we define a class of strategies which select actions depending on the common belief θ_t instead of the entire common information h_t^R .

Definition 2. *We define the set of common belief based strategies $\mathcal{G}^C \subset \mathcal{G}^L \times \mathcal{G}^R$. For any $(g^L, g^R) \in \mathcal{G}^C$ we have the following. At any time t , for each h_t^R , let θ_t be the common belief constructed by (13)-(15) in Lemma 3. Then, there exists $g_t^{R,C} : \Delta(\mathbb{R}^{n_x}) \mapsto \mathbb{R}^{n_R}$, $\bar{g}_t^{L,C} : \Delta(\mathbb{R}^{n_x}) \mapsto \mathbb{R}^{n_L}$ and $\tilde{g}_t^{L,C} : \Delta(\mathbb{R}^{n_x}) \mapsto \mathcal{Q}^{\theta_t}$ such that*

$$g_t^R(h_t^R) = g_t^{R,C}(\theta_t), \quad (17)$$

$$g_t^L(h_t^R) = \bar{g}_t^{L,C}(\theta_t) + \tilde{g}_t^{L,C}(\theta_t)(x_t). \quad (18)$$

Our main result of this section is the dynamic program provided in the theorem below.

Theorem 1. Suppose there are value functions $\{V_t : \Delta(\mathbb{R}^{n_x}) \mapsto \mathbb{R} \text{ for } t = 0, 1, \dots, T+1\}$ such that $V_{T+1} = 0$, and for each time t and for each $\theta_t \in \Delta(\mathbb{R}^{n_x})$

$$V_t(\theta_t) = \min_{q_t \in \mathcal{Q}^{\theta_t}} \left\{ \min_{\bar{u}_t^L \in \mathbb{R}^{n_L}, u_t^R \in \mathbb{R}^{n_R}} \left\{ \int_{\mathbb{R}^{n_x}} c_t(x_t, \bar{u}_t^L + q_t(x_t), u_t^R) \theta_t(dx_t) \right. \right. \\ \left. \left. + (1-p) \int_{\mathbb{R}^{n_x}} V_{t+1}(\varphi(x_{t+1})) \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)(dx_{t+1}) \right. \right. \\ \left. \left. + pV_{t+1}(\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)) \right\} \right\}. \quad (19)$$

If there are strategies $(g^{L*}, g^{R*}) \in \mathcal{G}^C$ with

$$g_t^{R*}(h_t^R) = g_t^{R,C*}(\theta_t), \quad (20)$$

$$g_t^{L*}(h_t^L) = \bar{g}_t^{L,C*}(\theta_t) + \tilde{g}_t^{L,C*}(\theta_t)(x_t) \quad (21)$$

such that for each h_t^R

$$u_t^{R*} = g_t^{R,C*}(\theta_t), \quad \bar{u}_t^{L*} = \bar{g}_t^{L,C*}(\theta_t), \quad q_t^* = \tilde{g}_t^{L,C*}(\theta_t), \quad (22)$$

achieve the minimum in the definition of $V_t(\theta_t)$, where θ_t is the common belief constructed by (13)-(15) in Lemma 3. Then g^{L*}, g^{R*} are optimal.

Theorem 1 provides a dynamic program to solve the two-controller problem. However, there are two challenges in solving the dynamic program. First, it is a dynamic program on the belief space $\Delta(\mathbb{R}^{n_x})$ which is infinite dimensional. Second, each step of the dynamic program involves a functional optimization over the functional space \mathcal{Q}^θ . Nevertheless, in the next section, we show that it is possible to find an exact solution to the dynamic program of Theorem 1, and provide optimal strategies for the controllers.

IV. OPTIMAL CONTROL STRATEGIES

In this section, we identify the structure of the value function in the dynamic program (19). Using the structure, we explicitly solve the dynamic program and obtain the optimal strategies for Problem 1.

For a vector x and a matrix G , we use

$$QF(G, x) = x^\top G x = \text{tr}(G x x^\top) \quad (23)$$

to denote the quadratic form.

The main result of this section, stated in the theorem below, presents the structure of the value function and an explicit optimal solution of the dynamic program (19).

Theorem 2. For any θ_t and any time t , the value function of the dynamic program (19) in Theorem 1 is given by

$$V_t(\theta_t) = QF(P_t, \mu(\theta_t)) + \text{tr}(\tilde{P}_t \text{cov}(\theta_t)) + e_t, \quad (24)$$

$$e_t = \sum_{s=t}^T \text{tr}(((1-p)P_{s+1} + p\tilde{P}_{s+1}) \text{cov}(\pi_{W_s})), \quad (25)$$

and the optimal solution is given by

$$\begin{bmatrix} \bar{u}_t^{L*} \\ u_t^{R*} \end{bmatrix} = \begin{bmatrix} \bar{g}_t^{L,C*}(\theta_t) \\ g_t^{R,C*}(\theta_t) \end{bmatrix} \\ = - \begin{bmatrix} G_t^{LL} & G_t^{LR} \\ G_t^{RL} & G_t^{RR} \end{bmatrix}^{-1} \begin{bmatrix} G_t^{LX} \\ G_t^{RX} \end{bmatrix} \mu(\theta_t), \quad (26)$$

$$q_t^*(x_t) = \tilde{g}_t^{L,C*}(\theta_t)(x_t) \\ = - (\tilde{G}_t^{LL})^{-1} \tilde{G}_t^{LX} (x_t - \mu(\theta_t)). \quad (27)$$

The matrices $P_t, G_t, H_t, \tilde{P}_t, \tilde{G}_t, \tilde{H}_t$ defined recursively below are symmetric positive semi-definite (PSD); G_t and \tilde{G}_t are symmetric positive definite (PD).

$$P_{T+1} = \tilde{P}_{T+1} = \mathbf{0}, \text{ the all zeros matrix}, \quad (28)$$

$$P_t = G_t^{XX}$$

$$- \begin{bmatrix} G_t^{XL} & G_t^{XR} \end{bmatrix} \begin{bmatrix} G_t^{LL} & G_t^{LR} \\ G_t^{RL} & G_t^{RR} \end{bmatrix}^{-1} \begin{bmatrix} G_t^{LX} \\ G_t^{RX} \end{bmatrix}, \quad (29)$$

$$G_t = \begin{bmatrix} G_t^{XX} & G_t^{XL} & G_t^{XR} \\ G_t^{LX} & G_t^{LL} & G_t^{LR} \\ G_t^{RX} & G_t^{RL} & G_t^{RR} \end{bmatrix} = R_t + H_t, \quad (30)$$

$$H_t = [A, B^L, B^R]^\top P_{t+1} [A, B^L, B^R], \quad (31)$$

$$\tilde{P}_t = \tilde{G}_t^{XX} - \tilde{G}_t^{XL} (\tilde{G}_t^{LL})^{-1} \tilde{G}_t^{LX}, \quad (32)$$

$$\tilde{G}_t = \begin{bmatrix} \tilde{G}_t^{XX} & \tilde{G}_t^{XL} & \tilde{G}_t^{XR} \\ \tilde{G}_t^{LX} & \tilde{G}_t^{LL} & \tilde{G}_t^{LR} \\ \tilde{G}_t^{RX} & \tilde{G}_t^{RL} & \tilde{G}_t^{RR} \end{bmatrix} \\ = R_t + (1-p)H_t + p\tilde{H}_t, \quad (33)$$

$$\tilde{H}_t = [A, B^L, B^R]^\top \tilde{P}_{t+1} [A, B^L, B^R]. \quad (34)$$

The proof of Theorem 2 relies on the following lemma for quadratic optimization problems.

Lemma 4. Let $G = \begin{bmatrix} G^{XX} & G^{XU} \\ G^{UX} & G^{UU} \end{bmatrix}$ be a PD matrix and $P := G^{XX} - G^{XU} (G^{UU})^{-1} G^{UX}$ be the Schur complement of G^{UU} of G .

(a) For any constant vector $x \in \mathbb{R}^n$,

$$\min_{u \in \mathbb{R}^m} QF(G, \text{vec}(x, u)) = QF(P, x) \quad (35)$$

with optimal solution

$$u^* = - (G^{UU})^{-1} G^{UX} x. \quad (36)$$

(b) For any $\theta \in \Delta(\mathbb{R}^n)$, let X^θ be a random variable with distribution θ , then

$$\min_{q \in \mathcal{Q}^\theta} \text{tr}(G \text{cov}(\text{vec}(X^\theta, q(X^\theta)))) \\ = \text{tr}(P \text{cov}(X^\theta)) \quad (37)$$

with optimal solution

$$q^*(X^\theta) = - (G^{UU})^{-1} G^{UX} (X^\theta - \mu(\theta)). \quad (38)$$

Using Lemma 4, we present a sketch of the proof of Theorem 2. The complete proof is in the Appendix.

Sketch of the proof of Theorem 2. The proof is done by induction. Suppose the result is true at $t + 1$, then at time t

- Show that G_t, \tilde{G}_t are PD.
- Apply the induction hypothesis for (24) and the sequential update of common belief in Lemma 3 to obtain

$$V_t(\theta_t) = \min_{q_t \in \mathcal{Q}^{\theta_t}} \left\{ \min_{\bar{u}_t^L \in \mathbb{R}^{n_L}, u_t^R \in \mathbb{R}^{n_R}} \left\{ QF(G_t, \mathbb{E}[S_t^{\theta_t}]) + \text{tr}(\tilde{G}_t \text{cov}(S_t^{\theta_t})) \right\} \right\}. \quad (39)$$

In the above equation, $S_t^{\theta_t} := \text{vec}(X^{\theta_t}, \bar{u}_t^L + q_t(X^{\theta_t}), u_t^R)$ where X^{θ_t} is a random vector with distribution θ_t .

- Since $q_t \in \mathcal{Q}^{\theta_t}$, $\mathbb{E}[q_t(X^{\theta_t})] = 0$ and consequently, $\mathbb{E}[S_t^{\theta_t}]$ depends only on u_t^R, \bar{u}_t^L . Furthermore, $\text{cov}(S_t^{\theta_t})$ depends only on q_t . Hence, (39) is equivalent to solving the following optimization problems

$$\min_{u_t^R, \bar{u}_t^L} QF(G_t, \text{vec}(\mathbb{E}[X^{\theta_t}], \bar{u}_t^L, u_t^R)), \quad (40)$$

$$\min_{q_t \in \mathcal{Q}^{\theta_t}} \text{tr}(\tilde{G}_t \text{cov}(\text{vec}(X^{\theta_t}, q_t(X^{\theta_t}), 0))). \quad (41)$$

- Apply Lemma 4 to problems (40) and (41), then we get (24) and the optimal solution at t . \square

From Theorem 1 and Theorem 2, we can explicitly compute the optimal strategies for Problem 1. The optimal strategies of controllers C^L and C^R are shown in the following theorem.

Theorem 3. *The optimal strategies of Problem 1 are given by*

$$\begin{bmatrix} \bar{U}_t^{L*} \\ U_t^{R*} \end{bmatrix} = - \begin{bmatrix} G_t^{LL} & G_t^{LR} \\ G_t^{RL} & G_t^{RR} \end{bmatrix}^{-1} \begin{bmatrix} G_t^{LX} \\ G_t^{RX} \end{bmatrix} \hat{X}_t, \quad (42)$$

$$U_t^{L*} = \bar{U}_t^{L*} - (\tilde{G}_t^{LL})^{-1} \tilde{G}_t^{LX} (X_t - \hat{X}_t), \quad (43)$$

where \hat{X}_t is the estimate (conditional expectation) of X_t based on the common information H_t^R . \hat{X}_t can be computed recursively according to

$$\hat{X}_0 = \begin{cases} \mu(\pi_{X_0}) & \text{if } Z_0 = \emptyset, \\ X_0 & \text{if } Z_0 = X_0. \end{cases} \quad (44)$$

$$\hat{X}_{t+1} = \begin{cases} A\hat{X}_t + B^L \bar{U}_t^{L*} + B^R U_t^{R*} & \text{if } Z_{t+1} = \emptyset, \\ X_{t+1} & \text{if } Z_{t+1} = X_{t+1}. \end{cases} \quad (45)$$

Theorem 3 shows that the optimal control strategy of C^R is linear in the estimated state \hat{X}_t , and the optimal control strategy of C^L is linear in both the actual state X_t and the estimated state \hat{X}_t . Note that even though the local controller C^L perfectly observes the system state, C^L still needs to compute the estimated state \hat{X}_t to make optimal decisions.

V. CONCLUSION

We considered a decentralized optimal control problem for a linear plant controlled by two controllers, a local controller and a remote controller. The local controller directly observes the state of the plant and can inform the remote controller of the plant state through a packet-drop channel with acknowledgments. We provided a dynamic program for this decentralized control problem using the common information approach. Although our problem is not partially nested, we obtained explicit optimal strategies for the two controllers. In the optimal strategies, both controllers compute a common estimate of the plant state based on the common information. The remote controller's action is linear in the common estimated state, and the local controller's action is linear in both the actual state and the common estimated state.

REFERENCES

- [1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, pp. 138–162, Jan 2007.
- [2] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, pp. 163–187, Jan 2007.
- [3] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.
- [4] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. S. Sastry, "Optimal control with unreliable communication: the TCP case," in *American Control Conference, 2005. Proceedings of the 2005*, pp. 3354–3359 vol. 5, June 2005.
- [5] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. Sastry, "Optimal linear LQG control over lossy networks without packet acknowledgment," in *Decision and Control, 2006 45th IEEE Conference on*, pp. 392–397, Dec 2006.
- [6] N. Elia and J. N. Eisenbeis, "Limitations of linear remote control over packet drop networks," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 5, pp. 5152–5157 Vol.5, Dec 2004.
- [7] E. Garone, B. Sinopoli, and A. Casavola, "LQG control over lossy TCP-like networks with probabilistic packet acknowledgements," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pp. 2686–2691, Dec 2008.
- [8] G. M. Lipsa and N. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Trans. Autom. Control*, vol. 56, pp. 2013–2025, Sept. 2011.
- [9] A. Nayyar, T. Basar, D. Teneketzis, and V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2246–2260, 2013.
- [10] R. Bansal and T. Başar, "Simultaneous design of measurement channels and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, pp. 679–694, 1989.
- [11] S. Tatikonda, A. Sahai, and S. K. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Autom. Control*, vol. 49, pp. 1549–1561, Sept. 2004.
- [12] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proc. IEEE*, vol. 95, pp. 108–137, Jan. 2007.
- [13] A. Molin and S. Hirche, "On the optimality of certainty equivalence for event-triggered control systems," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 470–474, 2013.
- [14] M. Rabi, C. Ramesh, and K. H. Johansson, "Separated design of encoder and controller for networked linear quadratic optimal control," *arXiv preprint arXiv:1405.0135*, 2014.
- [15] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal of Optimal Control*, vol. 6, no. 1, pp. 131–147, 1968.
- [16] G. M. Lipsa and N. C. Martins, "Optimal memoryless control in Gaussian noise: A simple counterexample," *Automatica*, vol. 47, no. 3, pp. 552–558, 2011.

- [17] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, vol. 36, no. 9, pp. 1249–1274, 2000.
- [18] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
- [19] R. Radner, "Team decision problems," *Annals of Mathematical Statistics*, vol. 33, pp. 857–881, 1962.
- [20] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems—Part I," *IEEE Trans. Autom. Control*, vol. 17, no. 1, pp. 15–22, 1972.
- [21] A. Lamperski and J. Doyle, "On the structure of state-feedback LQG controllers for distributed systems with communication delays," in *Proc. 50th IEEE Conf. Decision and Control and European Control Conf (CDC-ECC)*, pp. 6901–6906, Dec. 2011.
- [22] L. Lessard and A. Nayyar, "Structural results and explicit solution for two-player LQG systems on a finite time horizon," in *IEEE 52nd Annual Conference on Decision and Control (CDC)*, 2013, pp. 6542–6549, Dec 2013.
- [23] P. Shah and P. Parrilo, " \mathcal{H}_2 -optimal decentralized control over posets: A state-space solution for state-feedback," *IEEE Trans. Autom. Control*, vol. 58, pp. 3084–3096, Dec. 2013.
- [24] A. Nayyar and L. Lessard, "Structural results for partially nested lqg systems over graphs," in *American Control Conference (ACC)*, 2015, pp. 5457–5464, July 2015.
- [25] L. Lessard and S. Lall, "Optimal control of two-player systems with output feedback," *IEEE Transactions on Automatic Control*, vol. 60, pp. 2129–2144, Aug. 2015.
- [26] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Autom. Control*, pp. 2773–2786, Dec. 2009.
- [27] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 274–286, 2006.
- [28] S. M. Asghari and A. Nayyar, "Decentralized control problems with substitutable actions," in *IEEE 54th Annual Conference on Decision and Control (CDC)*, 2015, Dec. 2015.
- [29] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

APPENDIX

Proof of Lemma 1. Consider an arbitrary but fixed strategy g^R of C^R . Then the control problem of C^L becomes a MDP with state $\hat{H}_t^L = \text{vec}(X_t, H_t^R)$. From the theory of MDP we know that C^L can use only \hat{H}_t^L to make the decision at t without loss of optimality. \square

Proof of Lemma 3. At time $t = 0$, $h_0^R = z_0$. According to (7), for any $E \in \mathbb{R}^{n_x}$,

$$\begin{aligned} \theta_0(X_0 \in E) &= \mathbb{P}(X_0 \in E | z_0) = \mathbb{P}(X_0 \in E | Z_0 = z_0) = \\ &\begin{cases} \mathbb{P}(X_0 \in E | \Gamma_0 = 0) = \mathbb{P}(X_0 \in E) = \pi_{X_0}(E) & \text{if } z_0 = \emptyset, \\ \mathbb{P}(X_0 \in E | X_0 = x_0) = \varphi(x_0)(E) & \text{if } z_0 = x_0 \end{cases} \end{aligned}$$

which gives (13). At time $t + 1$, for any $E \in \mathbb{R}^{n_x}$, if $z_{t+1} = x_{t+1}$, then

$$\begin{aligned} \mathbb{P}^{g_{0:t}^L, g_{0:t}^R}(X_{t+1} \in E | h_{t+1}^R) &= \mathbb{P}(X_{t+1} \in E | x_{t+1}) \\ &= \mathbb{P}(X_{t+1} \in E | X_{t+1} = x_{t+1}) = \varphi(x_{t+1})(E). \end{aligned} \quad (46)$$

If $z_{t+1} = \emptyset$, then

$$\begin{aligned} \mathbb{P}^{g_{0:t}^L, g_{0:t}^R}(X_{t+1} \in E | h_{t+1}^R) &= \mathbb{P}^{g_{0:t}^L, g_{0:t}^R}(X_{t+1} \in E | h_t^R, \Gamma_{t+1} = 0) \\ &= \mathbb{P}^{g_{0:t}^L, g_{0:t}^R}(AX_t + B^L U_t^L + B^R U_t^R + W_t \in E | h_t^R) \\ &= \mathbb{P}(AX_t + B^L(\bar{u}_t^L + q_t(X_t)) + B^R u_t^R + W_t \in E | h_t^R) \\ &= \int_{\mathbb{R}^{n_x}} \int_{\mathbb{R}^{n_x}} \mathbb{1}_E(Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t) \\ &\quad \theta_t(dx_t) \pi_{W_t}(dw_t) \end{aligned} \quad (47)$$

where the third equality follows from Lemma 2. Furthermore, the last equality of (47) is true because W_t is independent of all previous random variables, and distribution of X_t given h_t^R is θ_t . Note that according to (7), (46), and (47), θ_{t+1} is only a function of $\theta_t, u_t^R, \bar{u}_t^L, q_t, z_{t+1}$. Hence, we can write it as $\theta_{t+1} = \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, z_{t+1})$ where $\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, x_{t+1})$ is given by (46) and $\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)$ is given by (47). \square

Proof of Theorem 1. Suppose the strategies $(g^{L*}, g^{R*}) \in \mathcal{G}^C$ satisfy (20)-(22). We prove by induction that for any $g^L \in \mathcal{G}^L, g^R \in \mathcal{G}^R, V_t(\mathbb{P}^{g_{0:t-1}^L, g_{0:t-1}^R}(dx_t | h_t^R))$ is a measurable function with respect to h_t^R , and for any information $h_t^R \in \mathcal{H}_t^R$ we have

$$\begin{aligned} &\mathbb{E}^{g_{0:t-1}^L, g_{0:t-1}^R, g_{t:T}^{L*}, g_{t:T}^{R*}} \left[\sum_{s=t}^T c_s(X_s, U_s^L, U_s^R) \middle| h_t^R \right] \\ &= V_t(\mathbb{P}^{g_{0:t-1}^L, g_{0:t-1}^R}(dx_t | h_t^R)) \end{aligned} \quad (48)$$

$$\leq \mathbb{E}^{g^L, g^R} \left[\sum_{s=t}^T c_s(X_s, U_s^L, U_s^R) \middle| h_t^R \right]. \quad (49)$$

Note that the above equation at $t = 0$ gives the optimality of g^{L*}, g^{R*} for Problem 1.

At $T + 1$ (48) and (49) are true (all terms are defined to be 0 at $T + 1$). Suppose (48) and (49) are true at $t + 1$.

Consider any $g^L \in \mathcal{G}^L, g^R \in \mathcal{G}^R$ and any information $h_t^R \in \mathcal{H}_t^R$ at time t . Let $\theta_t(dx_t) = \mathbb{P}^{g_{0:t-1}^L, g_{0:t-1}^R}(dx_t | h_t^R)$ be the common belief given h_t^R under strategies $g_{0:t-1}^L, g_{0:t-1}^R$.

We first consider (48). For notational simplicity let $g' = \{g_{0:t-1}^L, g_{0:t-1}^R, g_{t:T}^{L*}, g_{t:T}^{R*}\}$. Let $u_t^{R*}, \bar{u}_t^{L*}, q_t^*$ be the minimizers defined by (22) for θ_t . From the smoothing property of conditional expectation we have

$$\begin{aligned} &\mathbb{E}^{g'} \left[\sum_{s=t}^T c_s(X_s, U_s^L, U_s^R) \middle| h_t^R \right] \\ &= \mathbb{E}^{g'} \left[\mathbb{E}^{g'} \left[\sum_{s=t+1}^T c_s(X_s, U_s^L, U_s^R) \middle| H_{t+1}^R \right] \middle| h_t^R \right] \\ &\quad + \mathbb{E}^{g'} [c_t(X_t, U_t^L, U_t^R) | h_t^R]. \end{aligned} \quad (50)$$

From the induction hypothesis, $V_{t+1}(\mathbb{P}^{g'}(dx_{t+1} | h_{t+1}^R))$ is measurable with respect to h_{t+1}^R , and (48) holds at $t + 1$.

Therefore,

$$\begin{aligned}
& \mathbb{E}^{g'} \left[\sum_{s=t}^T c_t(X_s, U_s^L, U_s^R) \middle| h_t^R \right] \\
&= \mathbb{E}^{g'} \left[V_{t+1}(\mathbb{P}^{g'}(dx_{t+1} | H_{t+1}^R)) \middle| h_t^R \right] \\
&\quad + \mathbb{E}^{g'} [c_t(X_t, U_t^L, U_t^R) | h_t^R] \\
&= \mathbb{E}^{g'} [V_{t+1}(\psi_t(\theta_t, u_t^{R*}, \bar{u}_t^{L*}, q_t^*, Z_{t+1})) | h_t^R] \\
&\quad + \int_{\mathbb{R}^{n_X}} c_t(x_t, \bar{u}_t^{L*} + q_t^*(x_t), u_t^{R*}) \theta_t(dx_t). \tag{51}
\end{aligned}$$

Note that X_{t+1} is independent of Γ_{t+1} . Since $\mathbb{P}(\Gamma_{t+1} = 0) = 1 - \mathbb{P}(\Gamma_{t+1} = 1) = p$, the first term in (51) becomes

$$\begin{aligned}
& \mathbb{E}^{g'} [V_{t+1}(\psi_t(\theta_t, u_t^{R*}, \bar{u}_t^{L*}, q_t^*, Z_{t+1})) | h_t^R] \\
&= p \mathbb{E}^{g'} [V_{t+1}(\psi_t(\theta_t, u_t^{R*}, \bar{u}_t^{L*}, q_t^*, Z_{t+1})) | h_t^R, \Gamma_{t+1} = 0] \\
&\quad + (1-p) \mathbb{E}^{g'} [V_{t+1}(\psi_t(\theta_t, u_t^{R*}, \bar{u}_t^{L*}, q_t^*, Z_{t+1})) | h_t^R, \Gamma_{t+1} = 1] \\
&= p V_{t+1}(\alpha_t) + (1-p) \mathbb{E}^{g'} [V_{t+1}(\varphi(X_{t+1})) | h_t^R] \\
&= p V_{t+1}(\alpha_t) + (1-p) \mathbb{E}^{g'} [V_{t+1}(\varphi(X_{t+1})) | h_t^R, \Gamma_{t+1} = 0] \\
&= p V_{t+1}(\alpha_t) + (1-p) \int_{\mathbb{R}^{n_X}} V_{t+1}(\varphi(x_{t+1})) \alpha_t(dx_{t+1}) \tag{52}
\end{aligned}$$

where $\alpha_t := \psi_t(\theta_t, u_t^{R*}, \bar{u}_t^{L*}, q_t^*, \emptyset)$. The third equality in (52) is true because X_{t+1} is independent of Γ_{t+1} . The last equality in (52) follows from Lemma 3.

Combining (51) and (52) we get (48) from the definition of the value function (19). Moreover, since $g' = \{g_{0:t-1}^L, g_{0:t-1}^R, g_{t:T}^{L*}, g_{t:T}^{R*}\}$ are all measurable functions, $V_t(\mathbb{P}^{g_{0:t-1}^L, g_{0:t-1}^R}(dx_t | h_t^R))$ equals to the conditional expectation $\mathbb{E}^{g'} [\sum_{s=t}^T c_t(X_s, U_s^L, U_s^R) | h_t^R]$ which is measurable with respect to h_t^R .

Now let's consider (49). Let u_t^R, \bar{u}_t^L, q_t be the variables defined by (15) from h_t^R and g^L, g^R . Following an argument similar to that of (50)-(52), we get

$$\begin{aligned}
& \mathbb{E}^{g^L, g^R} \left[\sum_{s=t}^T c_s(X_s, U_s^L, U_s^R) \middle| h_t^R \right] \\
&\geq \int_{\mathbb{R}^{n_X}} c_t(x_t, \bar{u}_t^L + q_t(x_t), u_t^R) \theta_t(dx_t) \\
&\quad + (1-p) \int_{\mathbb{R}^{n_X}} V_{t+1}(\varphi(x_{t+1})) \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset) (dx_{t+1}) \\
&\quad + p V_{t+1}(\psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)) \geq V_t(\theta_t). \tag{53}
\end{aligned}$$

The last inequality in (53) follows from the definition of the value function (19). This completes the proof of the induction step, and the proof of the theorem. \square

Proof of Lemma 4. The proof of the first part of Lemma 4 is trivial.

Now let's consider the second part of Lemma 4, the functional optimization problem (37). From the property of

trace and covariance we have

$$\begin{aligned}
& \text{tr} (G \text{cov} (\text{vec} (X^\theta, q(X^\theta)))) \\
&= \mathbb{E} [QF (G, \text{vec} (X^\theta, q(X^\theta)) - \mathbb{E} [\text{vec} (X^\theta, q(X^\theta))])] \\
&= \mathbb{E} [QF (G, \text{vec} (X^\theta - \mu(\theta), q(X^\theta)))] \tag{54}
\end{aligned}$$

where the last equation in (54) holds because $\mathbb{E} [q(X^\theta)] = 0$. Since θ is the distribution of X^θ , we have

$$\begin{aligned}
& \mathbb{E} [QF (G, \text{vec} (X^\theta - \mu(\theta), q(X^\theta)))] \\
&= \int_{\mathbb{R}^n} QF (G, \text{vec} (y - \mu(\theta), q(y))) \theta(dy) \tag{55}
\end{aligned}$$

Note that the function inside the integral of (55) has the quadratic form of the optimization problem (35) with $x = y - \mu(\theta)$ and $u = q(y)$. From the results of the first part of Lemma 4, for any $y \in \mathbb{R}^n$ we have

$$\begin{aligned}
& QF (G, \text{vec} (y - \mu(\theta), q(y))) \\
&\geq QF (G, \text{vec} (y - \mu(\theta), q^*(y))) = QF (P, y - \mu(\theta))
\end{aligned}$$

where q^* is the function given by (38). It is clear that q_t^* is measurable. Furthermore,

$$\mathbb{E} [q^*(X^\theta)] = \int_{\mathbb{R}^n} -(G^{UU})^{-1} G^{UX} (x - \mu(\theta)) \theta(dx) = 0.$$

Consequently, $q^* \in \mathcal{Q}^\theta$. Then q^* is the optimal solution to problem (37), and the optimal value is given by

$$\begin{aligned}
& \int_{\mathbb{R}^n} QF (G, \text{vec} (y - \mu(\theta), q^*(y))) \theta(dy) \\
&= \int_{\mathbb{R}^n} QF (P, y - \mu(\theta)) \theta(dy) = \mathbb{E} [QF (P, X^\theta - \mu(\theta))] \\
&= \text{tr} (P \text{cov} (\theta)). \tag{56}
\end{aligned}$$

\square

Proof of Theorem 2. The proof is done by induction.

At $T+1$, (24) is true since $P_{T+1} = \tilde{P}_{T+1} = 0$. Suppose (24) is true at $t+1$ and the matrices are all PSD and G_{t+1}, \tilde{G}_{t+1} are PD.

At time t , since P_{t+1} and \tilde{P}_t are PSD, H_t and \tilde{H}_t are PSD. Since R_t is PD, $G_t = R_t + H_t$ and $\tilde{G}_t = R_t + (1-p)H_t + p\tilde{H}_t$ are also PD. Then P_t is PSD because P_t is the Schur complement of $\begin{bmatrix} G_t^{LL} & G_t^{LR} \\ G_t^{RL} & G_t^{RR} \end{bmatrix}$ of the matrix G_t . Similarly, \tilde{P}_t is PSD because \tilde{P}_t is the Schur complement of \tilde{G}_t^{LL} of the matrix $\begin{bmatrix} \tilde{G}_t^{XX} & \tilde{G}_t^{XL} \\ \tilde{G}_t^{LX} & \tilde{G}_t^{LL} \end{bmatrix}$.

Let's now compute the value function at t given by (19) in Theorem 1. For notational simplicity, let $\alpha_t = \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)$.

We first consider the second term of the value function in (19). From the induction hypothesis we have

$$\begin{aligned}
& (1-p) \int_{\mathbb{R}^{n_X}} V_{t+1}(\varphi(x_{t+1})) \alpha_t(dx_{t+1}) \\
&= (1-p) \int_{\mathbb{R}^{n_X}} QF (P_{t+1}, x_{t+1}) \alpha_t(dx_{t+1}) + (1-p)e_{t+1} \\
&= (1-p) QF (P_{t+1}, \mu(\alpha_t)) \\
&\quad + (1-p) \text{tr} (P_{t+1} \text{cov}(\alpha_t)) + (1-p)e_{t+1}. \tag{57}
\end{aligned}$$

The last equality in (57) follows from the property of covariance. Similarly, the last term of (19) becomes

$$p V_{t+1}(\alpha_t) = p QF(P_{t+1}, \mu(\alpha_t)) + p \text{tr}(\tilde{P}_{t+1} \text{cov}(\alpha_t)) + p e_{t+1}. \quad (58)$$

Let $S_t^{\theta_t} := \text{vec}(X^{\theta_t}, \bar{u}_t^L + q_t(X^{\theta_t}), u_t^R)$ where X^{θ_t} is a random vector with distribution θ_t such that X^{θ_t} and W_t are independent. Note that from (16) in Lemma 3

$$Y_t^{\theta_t} := [A, B^L, B^R] S_t^{\theta_t} + W_t = AX^{\theta_t} + B^L(\bar{u}_t^L + q_t(X^{\theta_t})) + B^R u_t^R + W_t \quad (59)$$

is a random vector with distribution α_t . Then, combining (57) and (58), the last two terms of the value function becomes

$$\begin{aligned} & QF(P_{t+1}, \mu(\alpha_t)) \\ & + \text{tr}(((1-p)P_{t+1} + p\tilde{P}_{t+1}) \text{cov}(\alpha_t)) + e_{t+1} \\ & = QF(P_{t+1}, \mathbb{E}[Y_t^{\theta_t}]) \\ & + \text{tr}(((1-p)P_{t+1} + p\tilde{P}_{t+1}) \text{cov}(Y_t^{\theta_t})) + e_{t+1} \\ & = QF(H_t, \mathbb{E}[S_t^{\theta_t}]) + \text{tr}(((1-p)H_t + p\tilde{H}_t) \text{cov}(S_t^{\theta_t})) \\ & + \text{tr}(((1-p)P_{t+1} + p\tilde{P}_{t+1}) \text{cov}(\pi_{W_t})) + e_{t+1} \\ & = QF(H_t, \mathbb{E}[S_t^{\theta_t}]) + \text{tr}(((1-p)H_t + p\tilde{H}_t) \text{cov}(S_t^{\theta_t})) \\ & + e_t. \end{aligned} \quad (60)$$

Using the random vector $S_t^{\theta_t}$, we can write the first term of the value function as

$$\begin{aligned} & \int_{\mathbb{R}^{n_X}} c_t(x_t, \bar{u}_t^L + q_t(x_t), u_t^R) \theta_t(dx_t) = \mathbb{E}[QF(R_t, S_t^{\theta_t})] \\ & = QF(R_t, \mathbb{E}[S_t^{\theta_t}]) + \text{tr}(R_t \text{cov}(S_t^{\theta_t})) \end{aligned} \quad (61)$$

Now putting (60) and (61) into (19) we get

$$V_t(\theta_t) = e_t + \min_{q_t \in \mathcal{Q}^{\theta_t}} \left\{ \min_{u_t^R, \bar{u}_t^L} \left\{ QF(G_t, \mathbb{E}[S_t^{\theta_t}]) + \text{tr}(\tilde{G}_t \text{cov}(S_t^{\theta_t})) \right\} \right\}. \quad (62)$$

Note that $\mathbb{E}[q_t(X^{\theta_t})] = 0$ since $q_t \in \mathcal{Q}^{\theta_t}$, and consequently, $\mathbb{E}[S_t^{\theta_t}] = \text{vec}(\mu(\theta_t), \bar{u}_t^L, u_t^R)$ depends only on u_t^R, \bar{u}_t^L . Furthermore, $\text{cov}(S_t^{\theta_t}) = \text{cov}(\text{vec}(X^{\theta_t}, q_t(X^{\theta_t}), 0))$ depends only on the choice of q_t . Consequently, the optimization problem in the (19) can be further simplified to be

$$\begin{aligned} V_t(\theta_t) &= e_t + \min_{u_t^R, \bar{u}_t^L} QF(G_t, \text{vec}(\mu(\theta_t), \bar{u}_t^L, u_t^R)) \\ &+ \min_{q_t \in \mathcal{Q}^{\theta_t}} \text{tr}(\tilde{G}_t \text{cov}(\text{vec}(X^{\theta_t}, q_t(X^{\theta_t}), 0))). \end{aligned} \quad (63)$$

Now we need to solve the two optimization problems

$$\min_{u_t^R, \bar{u}_t^L} QF(G_t, \text{vec}(\mu(\theta_t), \bar{u}_t^L, u_t^R)), \quad (64)$$

$$\min_{q_t \in \mathcal{Q}^{\theta_t}} \text{tr}(\tilde{G}_t \text{cov}(\text{vec}(X^{\theta_t}, q_t(X^{\theta_t}), 0))). \quad (65)$$

Since G_t is PD, it follows by Lemma 4 that the optimal solution of (64) is given by (26) and

$$\min_{u_t^R, \bar{u}_t^L} QF(G_t, \text{vec}(\mu(\theta_t), \bar{u}_t^L, u_t^R)) = QF(P_t, \mu(\theta_t)). \quad (66)$$

Similarly, since \tilde{G}_t is also PD, Lemma 4 implies that the optimal solution of (65) is given by (27) and

$$\begin{aligned} & \min_{q_t \in \mathcal{Q}^{\theta_t}} \text{tr}(\tilde{G}_t \text{cov}(\text{vec}(X^{\theta_t}, q_t(X^{\theta_t}), 0))) \\ &= \text{tr}(\tilde{P}_t \text{cov}(\theta_t)). \end{aligned} \quad (67)$$

Finally, substituting (66) and (67) into (63) we obtain the (24) at t . This completes the proof of the induction step and the proof of the theorem. \square

Proof of Theorem 3. Let \hat{X}_t be the estimate (conditional expectation) of X_t based on the common information H_t^R . Then, for any realization h_t^R of H_t^R , $\hat{x}_t = \mu(\theta_t)$. This together with Theorems 1 and 2 result in (42) and (43). To show (44) and (45), note that at time $t = 0$, for any realization h_t^R of H_t^R ,

$$\begin{aligned} \hat{x}_0 &= \mu(\theta_0) = \int_{\mathbb{R}^{n_X}} y \theta_0(dy) \\ &= \begin{cases} \int_{\mathbb{R}^{n_X}} y \pi_{X_0}(dy) = \mu(\pi_{X_0}) & \text{if } z_0 = \emptyset, \\ \int_{\mathbb{R}^{n_X}} y \varphi(x_0)(dy) = x_0 & \text{if } z_0 = x_0. \end{cases} \end{aligned} \quad (68)$$

Therefore, (44) is true. Furthermore, at time $t + 1$ and for any realization h_t^R of H_t^R ,

$$\hat{x}_{t+1} = \mu(\theta_{t+1}) = \int_{\mathbb{R}^{n_X}} y \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, z_{t+1})(dy).$$

If $z_{t+1} = x_{t+1}$, then $\hat{x}_{t+1} = \int_{\mathbb{R}^{n_X}} y \varphi(x_{t+1})(dy) = x_{t+1}$. If $z_{t+1} = \emptyset$, then,

$$\begin{aligned} \hat{x}_{t+1} &= \int_{\mathbb{R}^{n_X}} y \psi_t(\theta_t, u_t^R, \bar{u}_t^L, q_t, \emptyset)(dy) = \int_{\mathbb{R}^{n_X}} y \int_{\mathbb{R}^{n_X}} \int_{\mathbb{R}^{n_X}} \\ & \mathbb{1}_{\{y\}}(Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t) \\ & \theta_t(dx_t) \pi_{W_t}(dw_t) dy \\ &= \int_{\mathbb{R}^{n_X}} \int_{\mathbb{R}^{n_X}} (Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t) \\ & \theta_t(dx_t) \pi_{W_t}(dw_t) = A\hat{x}_t + B^L \bar{u}_t^L + B^R u_t^R. \end{aligned} \quad (69)$$

where the third equality is true because

$$\begin{aligned} & \int_{\mathbb{R}^{n_X}} y \mathbb{1}_{\{y\}}(Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t) dy \\ &= Ax_t + B^L(\bar{u}_t^L + q_t(x_t)) + B^R u_t^R + w_t. \end{aligned}$$

Furthermore, the last equality is true because $q_t \in \mathcal{Q}^{\theta}$ and W_t is a zero mean random vector. Therefore, (45) is true and the proof is complete. \square